



# Approche instrumentale et cognitive de la prosodie du discours en français

Cristel Portes

## ► To cite this version:

Cristel Portes. Approche instrumentale et cognitive de la prosodie du discours en français. Travaux Interdisciplinaires sur la Parole et le Langage, 2002, 21, pp.101-119. hal-00134184

**HAL Id: hal-00134184**

**<https://hal.science/hal-00134184>**

Submitted on 1 Mar 2007

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# **APPROCHE INSTRUMENTALE ET COGNITIVE DE LA PROSODIE DU DISCOURS EN FRANÇAIS**

Cristel Portes

## **Résumé**

*Notre travail concerne la mise en évidence, à partir d'une étude de cas sur la production d'un unique locuteur, des indices prosodiques contribuant à signaler la structuration hiérarchique du message transmis. Une réflexion approfondie sur les mécanismes cognitifs à l'œuvre dans la production et la perception du discours oral spontané nous a conduit à élaborer une tâche perceptive, destinée à recueillir la segmentation d'un extrait de débat radiophonique en unités discursives découpées par les sujets à partir d'informations essentiellement auditives. Le produit de cette segmentation a pu dès lors servir de base à une analyse prosodique qui nous a permis de définir quelques marqueurs prosodiques utilisés par notre locuteur pour structurer son message.*

Mots-clés : prosodie, discours, production et perception, parole spontanée, processus cognitifs, argumentation.

## **Abstract**

*The aim of this work is to find prosodic markers of the hierarchical structure of a speech message from a case study. What we know about cognitive mechanisms which guide spontaneous speech production and perception leads us to elaborate a perceptual experimentation. Its purpose is to obtain the segmentation of a radiophonic debate corpus into units of discourse given essentially by auditory analysis. The segmentation product may then be submitted to a prosodic analysis, which allows to define some prosodic markers used by the speaker to structure his message.*

Keywords : prosody, discourse, production and perception, spontaneous speech, cognitive processes, argumentation.

---

PORTES, Cristel, (2002), Approche instrumentale et cognitive de la prosodie du discours en français, *Travaux Interdisciplinaires du Laboratoire Parole et Langage*, vol. 21, p. 101-119.

## 1. Introduction

Parmi les nombreuses fonctions qu'assume la prosodie dans la communication orale et dont on trouvera un inventaire circonstancié dans (Di Cristo, 2000), la fonction « structurale », c'est-à-dire la contribution de la prosodie à la segmentation du discours en unités hiérarchisées, occupe une place centrale, très souvent évoquée, mais relativement peu étudiée dans ses mécanismes précis.

Afin de contribuer à une meilleure connaissance de ces mécanismes, ce travail se propose de repérer quelques phénomènes prosodiques à l'œuvre dans la structuration du discours en français et d'identifier leur rôle dans l'établissement d'unités discursives et de relations entre ces unités. En effet, la linéarisation qu'impose la matérialité sonore de la parole, où les informations ne peuvent être présentées que les unes après les autres selon le déroulement linéaire du temps, rend difficile la conservation et le transport de l'information structurale. Notre hypothèse est que les indices prosodiques apportent une aide cruciale au récepteur du message lorsqu'il reconstitue l'architecture discursive en vue de l'interprétation. Selon nous, la mise en forme prosodique capturerait et permettrait de restituer des informations concernant la constitution de *blocs informationnels* (voir (Vallduví, 1991) pour cette notion) et surtout leurs relations hiérarchiques.

Pour approcher cette question encore peu abordée dans son ensemble, nous avons élaboré et expérimenté une méthodologie spécifique dans notre DEA intitulé *Approche du rôle de la prosodie dans la structuration du discours oral en français* (2000).

C'est cette méthodologie que nous voudrions présenter ici, et en particulier les questions que cette réflexion a fait surgir à propos des mécanismes cognitifs à l'œuvre dans la production et la perception de la parole spontanée. Nous présenterons enfin les résultats de notre première exploration des indices prosodiques structurant le discours avant d'évoquer les perspectives que nous a ouvertes ce travail.

## 2. Des questions de méthodologie

Avant d'aborder les problématiques de cette recherche, nous allons préciser brièvement la façon dont nous définissons le discours et la prosodie dans la présente étude, tant est lourde de conséquences la teneur de ces choix théoriques.

### 2.1. Notions préliminaires

#### 2.1.1. Discours

Pour les sciences du langage, la notion de discours a pris corps à partir d'une interrogation sur l'usage du langage en situation de communication. Cet usage s'est avéré hautement influencé par

le *contexte*, notion qui recense aussi bien les contraintes de la situation de communication que l'influence du co-texte.

Au premier stade de notre recherche, nous avons retenu une définition étendue du discours qui correspond à la définition du dictionnaire le Petit Robert : « l'exercice de la faculté de langage » et, par extension métonymique, le produit de cette faculté : « énoncé linguistique observable (...), par opposition au système abstrait que constitue la langue » (Rey-Debove et Rey, 1993). Nous y retrouvons non seulement la définition précise de « l'observable » accessible à notre analyse de linguiste, mais aussi celle du principal matériau dont dispose le récepteur du message pour construire son interprétation. L'exposé de notre problématique nous donnera l'occasion d'expliquer pourquoi ce support de la réception s'est imposé comme notre objet d'analyse.

### 2.1.2. Prosodie

La prosodie est en quelque sorte la musique du langage. Elle concerne le rythme de la parole et sa mélodie. Sa description suppose l'observation des variations de paramètres physiques tels que la fréquence fondamentale (*Fo*), l'intensité, et la durée des événements linguistiques. Parmi les nombreuses fréquences émises par l'appareil vocal, la fréquence fondamentale est la composante la plus basse, celle qui a la même période que l'ensemble du signal acoustique. Ce paramètre définit *l'intonation*, qui rend compte des variations de hauteur de la voix. La durée des événements linguistiques, celle des syllabes par exemple, est également un élément essentiel de la description prosodique. Le rôle joué par l'intensité est moins bien connu, surtout en ce qui concerne la parole spontanée, car l'étude de ce paramètre impose de lourdes contraintes techniques, qui rendent difficile l'accès aux données.

L'analyse prosodique d'un corpus de parole spontanée, que nous présentons en section 5 de cet article, s'en tiendra donc à décrire les phénomènes intonatifs et temporels utilisés par le locuteur que nous avons choisi d'étudier pour mettre en forme la structure de son message.

## 2.2. Problématique

Trois types de recherches ont nourri notre réflexion sur la nature de la structure discursive dont nous voulions extraire les marqueurs prosodiques pertinents : les recherches psycholinguistiques sur la production et la perception de la parole, la réflexion théorique en prosodie et les modèles d'analyse du discours.

### 2.2.1. Modèles psycholinguistiques

Les recherches psycholinguistiques portant sur la production de la parole ont donné lieu à une synthèse remarquable dans l'ouvrage de W. Levelt intitulé *Speaking* (Levelt, 1989). Levelt postule que le discours est généré de façon séquentielle, chaque séquence étant élaborée pendant que la

précédente est prononcée. Dans une perspective analogue, Chafe (Chafe, 1988) considère la segmentation de surface comme le reflet de la planification cognitive. Ce sont précisément les unités issues d'une telle segmentation que la prosodie contribuerait selon nous à mettre en forme. Sur le versant des études perceptives le modèle de compétition de E. Bates et B. MacWhinney (Bates et MacWhinney, 1989), à l'option théorique inverse (connexionniste versus la conception modulaire de Levelt), postule le traitement par le système de décodage d'un ensemble d'indices hétérogènes et concurrents dont la synthèse permet à l'auditeur de choisir une interprétation. Rien n'est dit cependant sur le moyen pour l'analyste d'accéder aux « segments » produits et perçus. Que disent à ce propos le prosodiste et l'analyste de discours ?

#### 2.2.2. La question de la constituance prosodique

De leur côté, les prosodistes ont conçu des modèles de la *constituance prosodique*, dont l'objectif est de recenser les unités découpées par la prosodie dans le flux sonore issu de la production de parole. Nespor et Vogel (Nespor & Vogel, 1986) décrivent ainsi un ensemble d'unités prosodiques soumises à une hiérarchie stricte. S'inspirant de ces travaux mais rejetant l'hypothèse de la hiérarchie stricte, D. Hirst et A. Di Cristo (Hirst & Di Cristo, 1998) font reposer leur description de la prosodie du français sur trois unités essentielles : l'unité tonale, l'unité rythmique et l'unité intonative.

Les travaux sur la constituance définissent leurs unités selon des critères exclusivement prosodiques. L'un des choix méthodologiques que nous aurions pu faire consistait à comparer « en parallèle » les unités définies par l'analyse prosodique avec celles que décrivent les autres niveaux de l'analyse linguistique et en particulier la syntaxe et l'analyse de discours. Mais cette méthode nous a paru à la fois très coûteuse à mettre en pratique et trop peu fondée dans la mesure où elle nécessite des choix théoriques qu'il faut eux-mêmes expliciter pour chaque niveau.

#### 2.2.3. Les unités de l'analyse de discours

En effet, les analystes du discours cherchent de leur côté à définir les unités du discours : *paratons* pour G. Brown et G. Yule (Brown & Yule, 1983) *actes de langages* et *interventions* pour E. Roulet (Roulet *et al.*, 2001), *clauses* et *périodes* pour A. Berrendonner (Berrendonner, 1993).

Nous aurions pu choisir l'un ou l'autre de ces modèles pour observer les phénomènes prosodiques à l'échelle des unités qu'il définit. Mais chacun réalise un découpage du flux de parole en fonction de ses préoccupations particulières : Brown et Yule cherchent à définir l'unité textuelle dont la cohérence est assurée par la référence à l'unicité du *topique* (le topique correspond à peu près au thème de l'énoncé, à « ce dont on parle ») ; Roulet délimite ses *actes de langages* à partir des

connecteurs selon des préoccupations pragmatiques ; Berrendonner définit les *clauses* en référence à une conception particulière de la notion de représentation discursive.

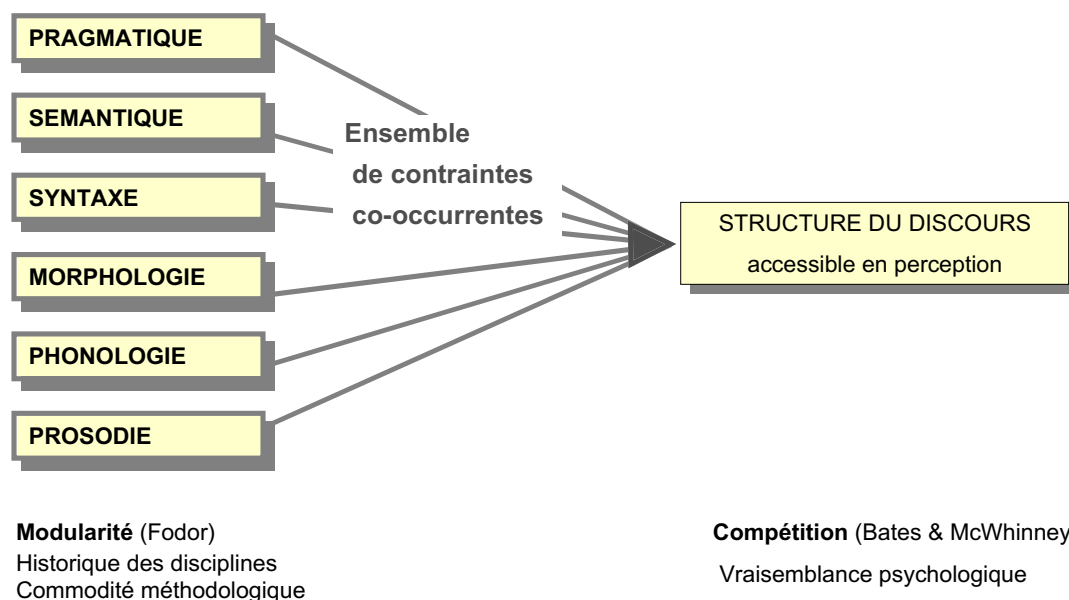
Compte tenu de la profusion des angles d'approche et des découpages auxquels ces conceptions, souvent antagonistes, donnent lieu, il nous a semblé impossible de choisir entre l'une ou l'autre de ces positions, toutes solidement étayées par des arguments convaincants, sans qu'aucune ne se révèle plus appropriée que les autres à nos besoins spécifiques.

Ce constat de l'impossible choix d'une segmentation prédéfinie par une théorie, qu'elle soit prosodique ou discursive, nous a conduit à expliciter les intuitions que nous avons au sujet des mécanismes cognitifs impliqués dans la production discursive.

### 3. Une réflexion sur les mécanismes cognitifs

En effet, nous pensons que le discours, et donc ses unités, est en fait le produit de l'ensemble des contraintes qui s'appliquent sur ces différents niveaux d'analyse que les linguistes se partagent : syntaxique, sémantique, pragmatique, prosodique, etc. Selon nous, l'observation d'un seul niveau ne saurait rendre compte *in fine* de ce qui préside au découpage du discours en *unités*, que seules les contraintes cognitives de la production et de la perception de la parole rendent indispensables (on trouvera du reste des objections faites à la notion d'unité pour la parole spontanée dans la thèse d'I. Guaitella (Guaitella, 1991)). L'hétérogénéité des découpages discursifs livrés par les différentes spécialités, et par les différentes écoles à l'intérieur d'une même spécialité, nous semble militer très fortement en faveur de notre conception.

Celle-ci a une autre conséquence si l'on pousse plus loin la réflexion. Les approches modulaires, en faveur desquelles argumenta J. Fodor (Fodor, 1983), et qu'ont adopté nombre de modèles comme celui d'E. Roulet (Roulet *et al.*, 2001), nous semblent appropriées pour mettre en œuvre une démarche descriptive claire des faits linguistiques. Leur valeur heuristique nous semble aussi indiscutable et précieuse. En revanche, des tentatives théoriques plus proches du connexionnisme, comme le « modèle de compétition » de Bates et MacWhinney (Bates et MacWhinney, 1989), nous paraissent présenter une plus grande vraisemblance psychologique. La figure 1 donne un aperçu synthétique de notre conception.



**Figure 1**  
*Arguments en faveur d'un découpage perceptif des unités de discours*

Seul le découpage *a posteriori*, réalisé pour les besoins de la perception, nous a paru pouvoir receler la trace de tous les niveaux de contrainte.

C'est pourquoi, afin d'obtenir des unités non tributaires d'un modèle choisi a priori, nous avons préféré recourir à une tâche perceptive. Cette démarche est aussi celle que préconisent les auteurs du projet PACOMUST, dont l'objectif est de définir un format permettant de constituer une base de données de parole multi-style ; voir (Astésano *et al.*, 1995).

La description de l'épreuve et de son exploitation fait l'objet de la section suivante.

#### 4. Une tâche perceptive pour une segmentation discursive

Comme la production de la parole contraint à linéariser des séquences discursives dans le temps, la perception doit reconstituer la structure du message à partir d'un continuum linéaire où sont stratifiées les informations provenant de tous les niveaux linguistiques. La première opération nécessaire à l'interprète consiste en un découpage du message en segments pertinents qu'il pourra

traiter successivement. Notre hypothèse était qu'une tâche devait nous permettre de restituer ce découpage.

Cependant, cette opération cognitive n'est pas d'ordinaire réalisée de façon consciente par l'interprète. Le découpage est fait « à la volée » au cours de la perception ; la conscience n'ayant besoin en définitive que du contenu restitué du message. La principale difficulté consistait donc à rendre compréhensible et réalisable une tâche ordinaire mais habituellement ignorée de ceux qui l'utilisent inconsciemment. Le problème s'est avéré d'autant plus délicat que nous présentions à la réception un corpus de parole spontanée dont la structure n'est pas immédiatement lisible comme l'est celle d'un texte écrit, ponctué et structuré en paragraphes. C'est pourquoi nous avons décidé d'utiliser les principales marques de la ponctuation : la virgule et le point comme marqueurs du découpage lors de l'épreuve. Familières à tous nos sujets, elles présentaient en outre l'avantage de rendre la consigne intuitivement compréhensible par tous, y compris appliquée à un corpus de parole spontanée.

#### 4.1. Protocole d'expérimentation

##### 4.1.1. Le corpus

Nous avons choisi un extrait d'émission radiophonique confrontant les arguments d'un avocat et d'un généticien sur la question de la preuve scientifique utilisée lors des procès. L'enregistrement radiophonique présente des qualités acoustiques propices à l'analyse prosodique. Le genre « débat » permet d'obtenir de la part des locuteurs des interventions suffisamment longues, et organisées par leur objectif argumentatif : la structure que nous cherchons à mettre en évidence s'y trouve nettement articulée.

Nous avons néanmoins sélectionné le locuteur le moins rompu à la maîtrise rhétorique (le généticien) pour garder la spontanéité d'un discours conçu en ligne.

L'extrait que nous avons choisi dure environ trois minutes. Il est constitué de trois interventions consécutives du généticien entrecoupées par des questions courtes du journaliste qui coordonne le débat.

##### 4.1.2. Les sujets

Deux populations de sujets se sont prêtées à l'expérimentation. Des sujets « naïfs » et des sujets « experts ».

Les « naïfs » : ayant suivi des études supérieures, ces sujets étaient suffisamment éduqués pour s'intéresser au débat (et donc faire l'effort cognitif d'interpréter le texte). En revanche, aucun



d'eux n'étant linguiste, aucun *a priori* métalinguistique lié à des options théoriques ne venait biaiser leurs choix structuraux.

Les « experts » : étaient des spécialistes de la prosodie. Nous attendions de leurs analyses qu'elles nous permettent d'affiner la structure issue des résultats de la première population.

#### 4.1.3. Les consignes

Au préalable, nous avons demandé à nos sujets d'écouter l'enregistrement de l'extrait choisi une première fois sans autre consigne que d'y prêter attention. Nous étions certaine, dès lors, qu'une première interprétation (et donc un premier découpage inconscient) était réalisé grâce à l'information auditive seule. En effet, la suite de la tâche réclamait un support écrit (la transcription orthographique du corpus) pour la notation des faits à identifier. Bien que nous insistions sur l'importance prédominante que les sujets devaient accorder à l'information auditive pour réaliser ce qui leur était demandé, il fallait contrecarrer l'influence forte exercée par l'habitude du maniement de l'écriture.

À l'issue de cette première écoute, nous demandions aux sujets de répondre aux consignes suivantes en s'appuyant sur les informations auditives en priorité.

Une première consigne les enjoignait à repérer les syllabes accentuées et à les souligner sur la transcription orthographique.

Ils devaient ensuite encadrer les expressions mises en évidence par une *emphase*, c'est-à-dire une insistance particulière.

Dans un troisième temps seulement nous leur demandions d'identifier à l'audition des blocs de parole (les unités que nous cherchions à mettre en évidence). Les limites de ces blocs ou « frontières » devaient être notées sur la transcription orthographique du corpus à l'aide de points et de virgules. Les points correspondaient à une frontière dite *conclusive* : qui marquait la fin d'un énoncé ; les virgules identifiaient les frontières dites *continuatives*, présentes entre des éléments constitutifs de la phrase mais laissant le message en suspens.

Enfin, il leur fallait associer à chaque frontière un coefficient de force de 1 à 5.

La figure 2 donne un aperçu du résultat obtenu à l'issue de la tâche :

alors je crois que les choses sont beaucoup plus compliquées que ça,<sup>1</sup> c'est-à-dire que lorsqu'on veut prouver la culpabilité d'un individu il faut apporter de **nombreux éléments**.<sup>3</sup> c'est vrai que la génétique a apporté,<sup>2</sup> un élément supplémentaire.<sup>3</sup> mais l'enquête est fait d'un ensemble de petits morceaux,<sup>2</sup> comme un puzzle,<sup>1</sup> qui va être petit à petit assemblé,<sup>2</sup> pour arriver à une,<sup>2</sup> **impression**,<sup>1</sup> **finale**,<sup>1</sup> qui sera ensuite présentée devant la cour d'assise. <sup>4</sup>

**Figure 2**  
*Exemple de résultat du recueil des réponses à la tâche*

Les syllabes soulignées sont celles sur lesquelles le sujet a identifié un accent ; les groupes de mots en caractères gras correspondent aux groupes emphatiques, les virgules et les points marquent respectivement les frontières *continuatives* et *conclusives*, et les nombres qui surmontent chaque signe de ponctuation représentent le degré de force assigné à la frontière concernée.

#### 4.1.4. Le traitement des données

Nous avons conservé les frontières (sans distinction de catégorie continuative ou conclusive), les accents et les groupes emphatiques détectés par au moins la moitié des sujets « naïfs » (six sur douze) ; pour les experts, trop peu nombreux (cinq), toutes les réponses ont été prises en compte. À chaque frontière a été affectée la moyenne des indices attribués à cette frontière.

#### 4.1.5. Résultats

##### • Résultats pour les naïfs

Deux faits intéressants ressortent de ces résultats.

Nous avons d'abord pu mettre en évidence que la totalité des frontières localisées par les sujets « naïfs » correspondaient à la détection d'une syllabe accentuée précédant immédiatement la frontière. Nous pouvions dès lors en conclure que, parmi les indices prosodiques pertinents pour ce découpage, la présence d'un accent jouait un rôle déterminant.

La seconde constatation concerne la distinction continuatif/conclusif établie par Delattre (Delattre, 1966a) et développée par Rossi (Rossi *et al.*, 1981). Les réponses des sujets ont montré un consensus perceptif beaucoup plus net sur les frontières conclusives (repérées par un point) que sur les frontières continuatives (identifiées par une virgule). Le tableau 1 donne des informations chiffrées à l'appui de cette constatation :

##### *Proportion relative des frontières continuatives et conclusives*

- 68 sites ponctués par 6/12 sujets ou plus
- 53 sites continuatifs (6 virgules ou plus)
- dont 8 (15%) sites ambigus (au moins 1 point)
- 15 sites conclusifs (6 points ou plus)

##### *Sites ambigus*

25% = 3 jugements conclusifs / 12 jugements de frontière  
 12,5% = 2 jugements conclusifs / 9 à 12 jugements de frontière  
 62,5% = 1 jugement conclusif / 9 à 12 jugements de frontière

##### *Sites conclusifs*

2/3 = plus de 9 jugements conclusifs / 9 à 12 jugements de frontière  
 Dont la moitié à 11 jugements conclusifs / 11 à 12 jugements de frontière

#### **Tableau 1**

*Importance numérique de la différence entre jugement de conclusivité (choix du point) et jugement de continuativité (choix de la virgule) apporté par les sujets naïfs*

Les frontières jugées continuatives le sont à l'unanimité dans 75% des cas : seuls 15% des cas sont ambigus. Et même alors, le jugement conclusif ne concerne jamais plus de trois sujets sur douze et le plus souvent un seul sujet sur douze. De même, les frontières jugées conclusives le sont deux fois sur trois par plus de neuf sujets sur douze. La différence paraît donc relativement claire et consensuelle pour nos sujets.

De plus la seule distinction significative qui émerge de la répartition des degrés de force de frontière concerne également la distinction, très nette, de ces deux types de frontières. En effet, 92% des frontières jugées continuatives sont affectées d'un coefficient de force de 1 ou 2. A l'inverse, la majorité des frontières jugées conclusives (67%), reçoit un coefficient de 4 ou 5.

Ces résultats renforcent l'idée d'une distinction binaire entre les fonctions conclusives et continuatives, hypothèse sur laquelle nous avons construit notre expérimentation. Une étude collective que nous avons menée ultérieurement, avec l'hypothèse initiale concurrente d'une échelle graduelle de continuation-conclusivité, donne elle aussi des résultats qui vont dans le sens d'une distinction catégorielle (Auran *et al.*, 2001).

- Résultats pour les experts

Les résultats obtenus auprès des experts ont confirmé les précédents : pour eux aussi, la perception d'une frontière est systématiquement corrélée à celle d'un accent, et la distinction continuatif-conclusif est bien définie. Mais surtout, les jugements de la population d'experts ont permis d'explicitier les quelques cas ambigus. Ils ont distingué deux types de frontières continuatives d'intensité différente, découpant le discours en unités de niveaux différents, les plus petites de l'ordre du syntagme, les plus grandes de l'ordre de la clause, les unes et les autres se combinant et s'agrégeant en un troisième groupe d'unités correspondant aux frontières conclusives.

Ainsi, et c'était là notre espoir, la tâche perceptive que nous venons de présenter a permis d'établir le découpage structurel de notre corpus. Nous pouvions dès lors soumettre les unités ainsi identifiées à l'analyse prosodique.

## 5. Quelques indices prosodiques qui signalent la structure du discours

La structure discursive établie à partir de la tâche perceptive que nous venons de décrire a donc pu servir de base à l'analyse prosodique que nous projetions. Pour la mener à bien, nous disposons des outils façonnés au laboratoire Parole et Langage et en particulier d'un modèle phonologique de l'intonation, élaboré par D. Hirst et A. Di Cristo (Hirst & Di Cristo, 1998), que nous voudrions présenter en quelques mots. En effet, ce modèle envisage la représentation cognitive de la prosodie qu'il décrit sur quatre niveaux :

- un niveau phonologique sous-jacent, où les matrices de tons L(low) et H(high) (voir les travaux de J. Pierrehumbert pour la conception originelle du modèle phonologique de l'intonation (Pierrehumbert, 1980)) définissent les unités intonatives (UI) et les unités tonales (UT), primitives de la constituance prosodique ;
- un niveau phonologique de surface, faisant également l'objet d'une représentation mentale, et attribuant à chaque proéminence une des valeurs de l'alphabet INTSINT (valeurs relatives : H=higher, L=lower, S= same) et valeurs absolues :T= top, B=bottom, M=mid) ;
- un niveau phonétique, où peut être modélisée la courbe de fréquence fondamentale ( $F_0$ ) débarrassée de ses variations non pertinentes pour la description intonative.
- un niveau physique enfin, rendant compte de la sortie acoustique du canal prosodique.

Le logiciel PHONEDIT développé au laboratoire Parole et Langage d'Aix-en-Provence permet le traitement du signal et sa représentation graphique en intégrant cette modélisation, notamment grâce à l'implémentation de l'algorithme MOMEL. Celui-ci permet la modélisation des courbes brutes de  $F_0$  en éliminant les variations micro-prosodiques issues de contraintes de bas niveau (notamment physiologiques de vibration des cordes vocales) et non pertinentes pour l'analyse de l'intonation. Ce modèle et ces outils ont servi de base à notre travail de description.

Celui-ci nous a permis de mettre en évidence (ou de retrouver) certains phénomènes prosodiques impliqués dans la structuration du discours spontané. Ces phénomènes se répartissent en deux grandes catégories selon qu'ils sont locaux et surviennent au niveau même de la frontière (ou de part et d'autre de cette frontière), ou globaux et concernent l'ensemble du domaine délimité par deux frontières.

## 5.1. Phénomènes prosodiques locaux

Nous inspirant des travaux de Delattre et de Rossi (cités plus haut), nous avons recensé différents types de patrons prosodiques précédant les frontières repérées par les virgules (patrons continuatifs) ou par les points (patrons conclusifs). Les pauses jouent aussi un rôle important au niveau local.

### 5.1.1. Patrons continuatifs

Examinant les courbes intonatives qui mettent en forme la fin des unités délimitées par une virgule, nous avons retrouvé dans notre corpus les deux types de patrons continuatifs définis par Delattre (Delattre, 1966a) : cet auteur distingue une continuation *simple* qui repère un « groupe à sens partiel » (qui ne peut être interprété qu'au regard de ce qui le suit), et une continuation *groupante* qui caractérise selon lui un « groupe à sens complet » (dont l'interprétation est possible, bien que d'autres informations soient attendues). Sur la figure 3, les premières sont notées par

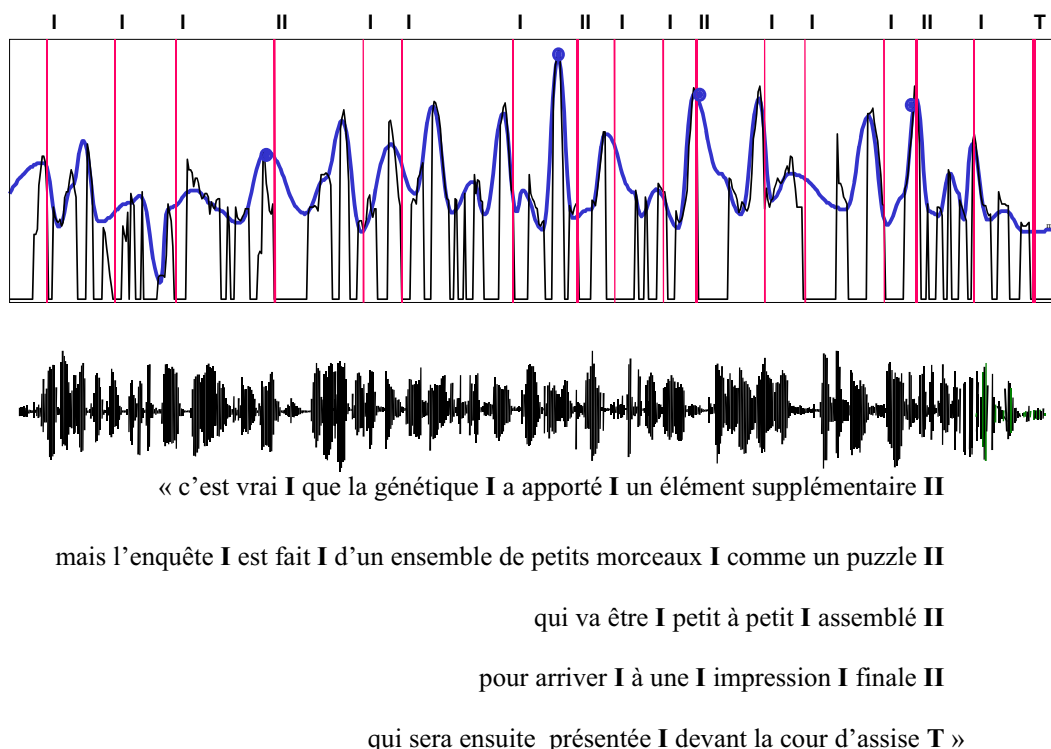
une barre verticale simple (I), les secondes par une barre double (II). Le marqueur T (pour « terminal ») signale une frontière conclusive.

c'est vrai I que la génétique I a apporté I un élément supplémentaire II mais l'enquête I est fait I d'un ensemble I de petits morceaux I comme un puzzle II qui va être I petit à petit I assemblé II pour arriver I à une I impression I finale II qui sera ensuite présentée I devant la cour d'assise T

### Figure 3

*Exemples d'unités à continuation simple(I) et à continuation groupante (II) dans un énoncé de notre locuteur*

Le contraste entre l'une et l'autre est marqué prosodiquement par la hauteur plus élevée atteinte par l'intonation sur la dernière syllabe et la pente plus verticale pour les continuations *groupantes*, traits que nous avons retrouvés de façon assez systématique, comme le montre la figure 4.



### Figure 4

*Contraste de hauteur et d'inclinaison de pente entre continuation simple (I) et continuation groupante (II)*

La figure 4 montre les variations de la courbe de fréquence fondamentale ( $F_0$ ) de l'énoncé présenté par la figure 3. La ligne courbe en trait épais représente la  $F_0$  modélisée par l'algorithme MOMEL qui n'en retient que les inflexions pertinentes au niveau intonatif (Hirst & Espesser, 1993). Les points signalent les pics de  $F_0$  correspondant aux frontières *groupantes* (II). On peut constater les valeurs globalement plus élevées de ces points par rapport aux autres pics de fréquence fondamentale et la pente plus raide de la montée qui les précède. L'écart entre le minima précédent et ces maxima de type II est aussi généralement plus important que pour les frontières *simples* (I).

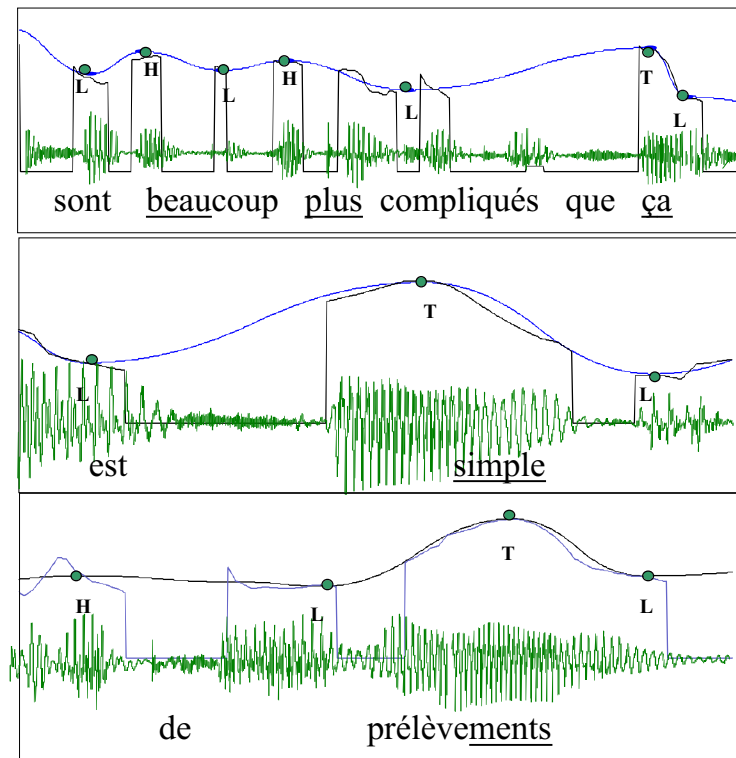
L'analyse prosodique des données issues du découpage opéré par notre sujet confirme donc la distinction établie par Delattre entre continuation simple et groupante. En revanche, nous ne le suivons pas dans son interprétation lorsqu'il identifie les segments isolés par la continuation simple à des « groupes à sens partiels », la continuation groupante délimitant selon ses dires des « groupes à sens plus complet ». Au-delà du peu de précision de ces formules, nous sommes amenés à constater plutôt l'hétérogénéité des segments ainsi distingués. Les frontières simples isolent en effet des fragments d'énoncés aussi différents qu'une relative à laquelle ne manque que son complément de lieu comme « qui sera ensuite présenté » (cf. figures 3 et 4 pour les exemples) et un adjectif séparé de son nom comme « finale ». Quant aux frontières groupantes, elles découpent des blocs qui peuvent correspondre à une clause entière comme « c'est vrai que la génétique a apporté un élément supplémentaire », mais tout aussi bien à un simple syntagme prépositionnel complément comme « pour arriver à une impression finale ». Selon nous, le jeu de ces deux types de frontières continuatives, associé à d'autres phénomènes prosodiques comme les variations du débit (que nous évoquerons plus loin) ou l'emphase, produit plutôt ce que l'on pourrait appeler des *effets de zoom*. On voit bien ici comment un découpage serré met en évidence « une I impression I finale » : expression qui a dans cet énoncé une haute valeur informationnelle. Au contraire, « qui sera ensuite présenté devant la cour d'assise » n'est pas mis en valeur par un découpage prosodique et le débit plus rapide de son énonciation indique selon nous que son contenu a moins d'importance pour le locuteur : l'information que ce segment véhicule se contente de rappeler une connaissance supposée partagée.

### 5.1.2. Patrons conclusifs

Les patrons intonatifs, qui correspondent aux frontières conclusives identifiées par nos sujets au moyen des points, ont eux aussi donné lieu à quelques observations intéressantes. Nous n'évoquerons pas ici les patrons intonatifs à pente simplement descendante, ni les patrons comportant une pente dont la descente vers le registre bas s'effectue en deux étapes (on parle

alors de downstep) : ce sont certes les patrons les plus représentés chez notre locuteur, mais la littérature les a souvent décrits.

Nous préférons parler d'un patron intonatif que nous avons appelé « circonflexe » ou « à inflexion ». Nous avons rencontré ce type de patron au niveau de frontières que seuls les sujets experts avaient identifiées comme terminales. La forme prosodique de ces patrons, représentée sur la figure 5, explique cette ambiguïté.



**Figure 5**

*Courbe de Fo modélisée et codage INTSINT des patrons conclusifs circonflexes de notre corpus*

Sur la dernière syllabe qui précède la frontière conclusive, l'intonation monte haut dans le registre du locuteur pour redescendre immédiatement et sur la même syllabe à une valeur beaucoup plus basse. Les valeurs Top (T) et Lower (L)) de l'alphabet INTSINT se succèdent ainsi sur la même

syllabe, créant un effet sonore très spécifique que seuls les experts ont interprété comme une clôture. Les frontières concernées correspondent dans tous les cas à une fin de proposition. De plus, les propositions ainsi terminées se trouvent placées en début de paragraphe, leur contenu présentant ce que les propositions suivantes développent. Il semble donc que notre locuteur utilise ce type de patron intonatif pour signaler ses énoncés introductifs.

Ce schème intonatif spécifique a été décrit par Delattre (Delattre, 1966b) sous le nom d'*intonation d'implication*. Il sera l'objet principal de notre thèse car nous avons pu constater son importance, en particulier pour l'argumentation, sur de larges corpus de débat.

### 5.1.3. Pauses

L'intonation n'est pas le seul phénomène prosodique à jouer un rôle dans la structuration informationnelle du discours. Les locuteurs utilisent aussi la souplesse de l'organisation temporelle et en particulier les pauses. Celles-ci ne sont pas systématiquement présentes au niveau des frontières terminales et leur durée y est en outre très variable (de 153,9 à 931,8 millisecondes). Elles présentent en revanche les durées les plus importantes au niveau des frontières que nous avons identifiées comme étant celles de *paragraphes* regroupant plusieurs segments conclusifs. Ces paragraphes ou paratons (selon la terminologie de Brown et Yule (Brown et Yule, 1983)) sont également signalés par des indices plus globaux.

## 5.2. Phénomènes prosodiques globaux

La mise en forme prosodique de la structure discursive ne concerne pas uniquement le niveau local de la frontière elle-même : certains phénomènes couvrent des domaines étendus, de taille variable. Trois d'entre eux jouent un rôle dans l'organisation discursive à l'échelle de notre corpus : les variations du débit, celles du registre, et des phénomènes de variation de la pente intonative.

### 5.2.1. Débit

Reprenons l'énoncé que retranscrit la figure 3 ci-dessus. Nous avons vu plus haut que le segment discursif « pour arriver à une impression finale » compris entre deux frontières groupantes (II) faisait l'objet d'un découpage très serré en segments simples (I) à l'origine de ce que nous avons appelé un « effet de zoom », grossissant dans ce cas, qui mettait en valeur le contenu du syntagme prépositionnel. Ce découpage étroit produit secondairement un ralentissement du débit lié à la multiplication des syllabes accentuées et porteuses d'un allongement, et à la présence de pauses de part et d'autre du mot « impression ». Le tempo accéléré du segment suivant « qui sera ensuite présentée devant la cour d'assise », renforce l'effet de mise en valeur du segment ralenti et signale du même coup la valeur informationnelle moindre de la relative.



De fait, notre locuteur utilise l'accélération essentiellement pour mettre en forme des énoncés à valeur informationnelle faible, par exemple à valeur illustrative. Mais nous trouvons aussi dans son discours un segment où l'accélération se superpose à l'emphase contribuant alors à la promotion informationnelle de l'énoncé. Nous constatons ici l'ambiguïté fonctionnelle de la prosodie : un même indice prenant des valeurs fonctionnelles différentes selon les cas. Cette complexité rend l'analyse prosodique plus apte à faire émerger des régularités qu'à énoncer des règles.

### 5.2.2. Registre

La plupart des études phonétiques ne mentionnent le registre que dans le cadre des différences entre locuteurs : en effet, chacun a une hauteur de voix différente, et l'étendue des hauteurs accessibles varie selon les individus. Cependant, il est possible d'observer des variations du registre à l'échelle d'un seul locuteur et notre analyse en donne deux exemples intéressants.

En premier lieu, notre locuteur utilise les variations de la hauteur de sa voix pour signaler la présence d'un énoncé parenthétique à valeur d'incise. La figure 6 retranscrit l'énoncé porteur de cet effet.

eh bien euh l'A.D.N. est cette molécule qui est le support de l'hérédité chez l'individu *euh bien sûr elle est connue du grand public parce que ce sont des anomalies de cette molécule qui vont entraîner les grandes maladies génétiques* mais elle nous intéresse nous surtout parce que cet A.D.N. nous a apporté des régions de variabilité qui vont faire que chaque individu va être différent de son voisin

**Figure 6**

*Énoncé parenthétique signalé par un abaissement du registre et son contexte*

La parenthèse, transcrite en italiques sur la figure 6, est produite dans un registre plus grave que les énoncés qui l'encadrent : la fréquence fondamentale ne dépasse pas le registre *infra-aigu* du locuteur (selon la définition du registre proposée par Rossi (Rossi, 1999)) sauf pour la dernière syllabe, porteuse de l'accent qui définit le patron continuatif groupant de fin d'unité intonative (voir l'UI dans le modèle de Hirst et Di Cristo cité plus haut) et de fin de parenthèse.

La deuxième fonction que notre locuteur assigne au contraste de registre est de contribuer à l'individualisation des paragraphes discursifs. Ainsi, dans l'extrait que nous avons choisi, la cohérence qui unit deux phrases en un paragraphe commun est renforcée par leur registre d'énonciation, plus grave que celui des paragraphes adjacents.

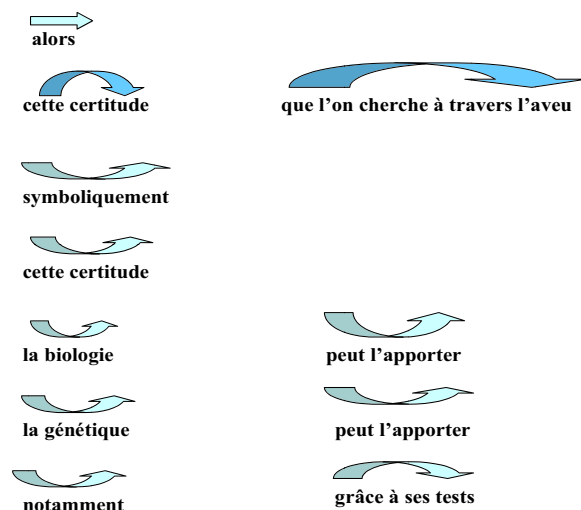
Là encore, nous constatons la pluri-fonctionnalité et le fonctionnement contrastif des indices prosodiques de structuration du discours.

### 5.2.3. Variations de la pente intonative

Contrairement à nos attentes, nous n'avons pas retrouvé chez notre locuteur les phénomènes d'abaissement que décrit la littérature et en particulier le *downdrift* : abaissement progressif des pics de fréquence fondamentale contrôlé par le locuteur (contrairement à la déclinaison dont l'origine est physiologique) sur le décours d'une unité signifiante.

En revanche, l'examen des variations de la pente intonative nous a permis de mettre en évidence un autre effet intéressant dont l'exemple le plus net nous est donné dans une intervention du journaliste qui modérat le débat radiophonique dont nous avons tiré notre corpus. Alors que l'intonation prototypique de la continuation est montante (Delattre, 1966a), deux unités consécutives dans la première question du journaliste sont mises en forme par une intonation globalement descendante comme le montre la figure 7.

Ce contraste de pente permet au journaliste, un professionnel de la parole rompu aux pratiques rhétoriques, de distinguer l'exposition du thème de son développement. La reprise lexicale du thème « cette certitude » renforce l'articulation rhétorique de son énoncé en deux mouvements clairement distincts.



**Figure 7**  
*Contraste de pente dans une intervention du journaliste*

## 6. Conclusion

La linéarisation qu'impose le flux de parole au message produit contraint fortement la transmission de la structure de ce message. Or, cette structure présente une grande complexité : comme nous avons essayé de le montrer, elle est hautement hiérarchisée, intégrant plusieurs niveaux d'organisation (phonologique, morphologique, syntaxique, sémantique, pragmatique, prosodique), sans compter la stratification que chaque niveau lui-même peut présenter (voir le modèle prosodique de Hirst et Di Cristo). Nous pensons que la prosodie joue un rôle important dans la conservation et le transport de cette information hiérarchique, structurale, que la linéarisation de la parole tend à « écraser ». Comme le peintre utilise les contrastes de luminosité pour donner au tableau la dimension de profondeur (technique de la perspective), le locuteur utilise en particulier les contrastes prosodiques pour encoder la structure pluridimensionnelle du discours. Mais c'est par l'intermédiaire de la perception, qui doit reconstituer le message à partir des indices transmis par l'onde acoustique, que l'analyste peut à son tour reconstituer la structure discursive et observer la contribution de la prosodie à sa transmission : la tâche perceptive que nous avons élaborée et mise en œuvre nous a permis d'accéder à une segmentation discursive valide perceptivement à partir de laquelle nous avons pu identifier et décrire quelques phénomènes prosodiques pertinents pour la mise en forme de la structure discursive.

Notre étude reste une étude de cas et il importe non seulement de vérifier dans d'autres corpus si les indices que nous avons décrits sont utilisés de la même façon par d'autres locuteurs, mais aussi de tester psycholinguistiquement de façon systématique la validité et la qualité perceptive de tels marqueurs structuraux. Par ailleurs, l'observation des relations rhétoriques entre les segments de discours que nous avons décrits, couplée à une analyse prosodique plus détaillée, devrait nous permettre d'affiner notre description de la contribution de la prosodie à l'organisation discursive. Tels sont quelques-uns des prolongements que nous aimerions donner à ce travail.

Ainsi, bien que notre recherche relève essentiellement d'une linguistique descriptive, nous espérons avoir montré que ses hypothèses et ses méthodes nécessitent une réflexion approfondie sur les mécanismes cognitifs à l'œuvre dans la production et la perception de la parole, et que ses résultats apportent des données nouvelles à la connaissance de ces mécanismes.

## Références bibliographiques

- ASTESANO, C. *et al.* (1995). Le projet PACOMUST, un corpus de PARole CONTinue MULTi-Styles : objectifs et choix méthodologiques. *Travaux de l'Institut de Phonétique d'Aix*, 16, p. 9-38.
- AURAN, C. *et al.* (2001), La distinction entre frontières conclusives et continuatives est-elle pertinente dans le discours spontané ?, *Actes des Journées Prosodie 2001*, Grenoble, France, 10-11 octobre 2001 (à paraître).
- BATES, E. & McWHINNEY, B. (1989). Functionnalism and the competition model. In McWhinney, B. & Bates, E. (eds). *The Crosslinguistic Study of Sentence Processing*. Cambridge University Press, Cambridge (UK), p. 3-73.
- BERRENDONNER, A. (1993). Périodes. In Parret, H. (ed.). *Temps et discours*. Presses universitaires de Louvain, Louvain, p. 47-61
- BROWN G. & YULE G. (1983). *Discourse Analysis*. Cambridge University Press. Cambridge, U.K.
- CHAFE, W. (1988) Linking Intonation Units. In Haiman, J. and Thomson, S.A. (eds), *Clause Combining in Grammar and Discourse*. John Benjamin, Amsterdam, p. 1-127
- DELATTRE, P. (1966a). La leçon d'intonation de Simone de Beauvoir. In Delattre (1966). *Studies in French and Comparative Phonetics*. Mouton & Co. La Hague, p. 75-82.
- DELATTRE, P. (1966b). Les dix intonations de base du français. *The French Review*, 40(1) : 1-14.
- DI CRISTO, A. (2000). Interpréter la prosodie. *Actes des XXIIIèmes Journées d'Etude sur la Parole*. Aussois, juin 2000, p. 13-29.
- FODOR, J. A. (1983). *The Modularity of Mind*. MIT Press, Cambridge, USA.
- GUAÏTELLA, I. (1991). *Rythme et parole ; comparaison critique du rythme de la lecture oralisée et de la parole spontanée*, Thèse de Doctorat, Université de Provence.
- HIRST, D.J. & ESPESSER, R. (1993). Automatic modelling of fundamental frequency using a quadratic spline function. *Travaux de l'Institut de Phonétique d'Aix*, 15, p. 71-85.
- HIRST, D. & DI CRISTO, A. (1998). *Intonation Systems, a Survey of Twenty Languages*. Cambridge University Press, Cambridge, U.K.
- LEVELT, J.M. (1989). *Speaking, from Intention to Articulation*. MIT Press, Cambridge, USA.
- NESPOR, M., VOGEL, I. (1986). *Prosodic Phonology*, Foris, Dordrecht.
- PIERREHUMBERT, J. (1980). *The Phonology and Phonetics of English Intonation*. PhD Thesis. MIT.
- REY-DEBOVE, J. & REY, A. (dir.) (1993). *Le nouveau petit Robert, dictionnaire alphabétique et analogique de la langue française*, Dictionnaires Le Robert, Paris.
- ROSSI, M. *et al.*, (1981). *L'intonation : De l'acoustique à la sémantique*, Klincksieck, Paris.
- ROSSI, M. (1999). *L'intonation, le système du français*, Ophrys, Paris.
- ROULET, E. *et al.* (2001). *Un modèle et un instrument d'analyse de l'organisation du discours*, Peter Lang, Berne.
- VALLDUVI, E. (1991) *The Informational Component*, Garland, New York.